RESEARCH ARTICLE

The geopolitics of vaccine media representation in Orbán's Hungary—an AI-supported sentiment analysis

Miklós Sebők[1] · Orsolya Ring[1] · Márk György Kis[1] ·
Martin Balázs Bánóczy[1] · Ágnes Dinnyés[1]

Abstract

Extant studies on the European media coverage of the COVID-19 pandemic generally posit a linear relationship between the severity of the public health emergency and the volume of media reports. However, domestic politics and geopolitics may also impact the saliency, distribution, and sentiment of coverage in different outlets. Under Viktor Orbán's illiberal leadership, Hungary sought deals for ventilators and vaccines from China and Russia—a deviation from joint European procurements. In this article, we conduct a content analysis of pro-government and Orbán-critical media to examine differences in their treatment of Eastern and Western vaccines. We relied on state-of-the-art deep learning analysis (a branch of articifial intelligence) to investigate all COVID-19-related articles (N = 72,339) published on three major Hungarian news portals between March 2020 and March 2022. We used a new fine-tuned BERT model for emotion analysis, the categories of which have been aggregated into three sentiment labels (positive, negative, and neutral). Our sentiment analysis results show a positive correlation between the number of sentences mentioning at least one of the vaccines and the (first) shots administered for only one outlet out of three. The pro-government portal in the sample showed more positivity towards Western vaccines than a hard-right, anti-government one. This latter also produced more positive reports concerning the Russian vaccine. These results shed light on the complex geopolitics of vaccine narratives in Hungarian media. Our research contributes not only to our understanding of illiberal media systems but also by sharing a new public dataset and a fine-tuned large language model that is applicable to alternative research questions and designs.

Keywords COVID-19 · Media representation · Geopolitics · Sentiment analysis · Emotion analysis · Large language models

---

Extended author information available on the last page of the article

🙋 Springer

Introduction

In April 2012, the Hungarian government approved a new economic strategy. One of its key objectives was a geographical diversification of exports and incoming foreign direct investment in the "Eastern opening" framework. The strategy was specifically aimed at strengthening Chinese and Russian relations [1]. Yet, according to international and domestic opinion polls, Hungarian citizens continued to show a pro-Western attitude, which was reflected in a favorable assessment of Western countries and institutions.[1]

That said, in the Orbán era, which started in 2010, almost half of the population took an "intermediate" position on West versus East. This was reinforced by systematic government campaigns against "Brussels" and Western liberal values (such as minority rights; see [2]). In addition, the popular reception of Russia (until at least the Ukraine war started in 2022) had undergone a significant positive change by the early 2020s, once more propped up by government communication via state- and party-controlled media [3, 4]. This suggests that public opinion on this issue was fundamentally shaped and dominated by political discourse in this period [4–6].

Hungary, led by Viktor Orbán, the illiberal right-wing populist prime minister, had long been an outlier among European Union countries regarding policy responses to various crises, the Russian invasion of Ukraine being one notable example [7]. COVID-19 response was no exception: the Orbán-government looked for solutions circumventing joint EU procurements and sought deals with China and Russia for both ventilators and vaccines [8–10].

As Griffen [11] argues, the Hungarian government led by Viktor Orbán is systematically eroding media freedom, pluralism, and independence. It has achieved this through measures such as undermining the independence of media watchdogs, manipulating access to sources, and controlling the information environment [4]. The rise of nationalist sentiment has further increased distrust and restrictions on the media [12].

During the first wave of the COVID-19 pandemic in 2020, the Hungarian government increased its control over official information to curb misinformation and rumors. However, media professionals, watchdogs, and politicians in Hungary and internationally criticized this approach for negatively affecting the right to information and freedom of expression [13, 14].

In this article, we investigate whether the opinions expressed on Eastern- and Western-produced vaccines can be linked to the geopolitical preferences of news outlets and their preferred political parties. Did Hungarian state-controlled media extol the virtues of Eastern vaccines to convince people to choose them? Can we detect meaningful differences in COVID-19 vaccine content between outlets with different ideological orientations and party preferences?

We studied these research questions using a state-of-the-art, deep learning-based large language model, a branch of the field of articifial intelligence. For the

---

[1] Hungarian public opinion on China in the age of COVID-19: Enthusiasm overshadowed by skepticism? https://ceias.eu/wp-content/uploads/2021/08/HU-poll-report.pdf; https://nezopont.hu/a-magyarok-meg-a-kozep-europai-atlagnal-is-amerika-partibbak/; https://merce.hu/2021/01/20/ki-szereti-kinat/.

annotation, we adopted a conventional codebook developed by Plutchik and Kellerman [15]. We aggregated its eight original emotions (not containing "neutral") into two sentiment categories: positive ("Joy", "Surprise", "Trust"), negative ("Anger", "Disgust", "Fear", "Sadness"), and we added a neutral category too (the Plutchik's "Anticipation" category, was not included in our manually coded data). This approach allowed us to carry out the sentiment analysis using a more refined emotion model.

This model was fine-tuned for emotion analysis with manually labeled Hungarian media data. We analyzed all COVID-19 vaccine-related articles (N=72 339) published on three major news portals between March 2020 and March 2022. Our content analysis focused on three groups of vaccines: Western (Pfizer/Comirnaty, AstraZeneca, Janssen, Moderna), Chinese (Sinopharm), and Russian (Sputnik V), where the latter two groups were also combined to form the Eastern vaccine group.

Our analysis shows a positive correlation between the number of vaccine-mentioning sentences and the number of first shots administered only for Origo, the largest pro-government outlet. Results also show significant differences between the outlets in attention paid to vaccines and vaccine group preferences. While the antigovernment, hard-right Alfahir favored the Chinese vaccine, Origo was very close to the opposition, liberal Telex, in its attitudes towards Western and Russian vaccines. Besides these empirical results for the interplay of the pandemic and geopolitics in an illiberal political-media system, we also contribute by sharing publicly our database, as well as our trained model (both of which can be applied to alternative research questions and designs).

In this article, we first provide an overview of the history of the COVID-19 response in Hungary. Next, we review the extant literature related to the (geo)politics and political communication of COVID-19 vaccination and present our hypotheses. This is followed by an introduction to our dataset and the methods applied. The Results section presents our key findings regarding media coverage of Western and Eastern vaccines, which we discuss in the next part. The Conclusion reviews the research design results, situates them within the literature, and contemplates avenues for future research.

News media, domestic politics, geopolitics

The impact of the COVID-19 pandemic on media coverage and geopolitical dynamics is analyzed by a fast-emerging literature. Mayer's study [16] highlights the role of the media in shaping public perceptions of COVID-19 hotspots such as Ischgl in Austria and the subsequent geopolitical blame game. Zunino [17] discusses the role of the media in information dissemination and the media's contribution to overinformation, polarisation, and misinformation. Sidarta [18] further examines how governments use the media to communicate policies and information related to COVID-19. Jalil et al. [19] and Ruf et al. [20] highlight the impact of geopolitical factors on news coverage, by focusing on social conflict and on political relations, respectively. Golan et al. [21] further examine the impact of domestic and national party identities on perceived media bias in foreign policy coverage. Finally, Karameti and

Bilalli-Zendeli [22] discuss the media's focus on the impact of the pandemic on the environment and its impact on political agendas.

Based on extant studies exploring exogenous factors affecting policy dynamics globally [23] and in Hungary [24, 25], we can assume that the media coverage of policy topics shows a punctuatuations are often exogenous. In the case of strong exogenous effects (as happened with the COVID-19 pandemic), the theoretical expectation is that all media outlets, regardless of ideological leaning, will adjust their coverage and feature the pandemic and vaccination as a key topic in this time frame.

Many of the studies that discuss the influence of media on public health narratives suggest that media coverage can precede and potentially steer public health interventions like vaccination programs [26–30]. Online news portals play an important role in shaping public opinion and conveying the views of the political actors they support, especially in Central-Eastern Europe, where such connections are traditionally strong. We, therefore, consider them particularly well-suited to shed light on the discourse surrounding government vaccination policies during the COVID-19 pandemic [9, 31–35]. As Papp and Nkansah [34] highlighted, in Hungary there was a strong political component motivating vaccine choice.

Geopolitics and the vaccine procurement of the Orbán government

The onset of the pandemic resulted in a slew of policy responses from the Hungarian government led by Viktor Orbán. The so-called Operational Body was established in late January 2020 to respond to the emergency, and it set up a website to inform citizens and encourage them to register for vaccination. However, the data provided by the government was criticized for being misleading [36]. The introduction of legislation granted Orbán emergency powers, including the ability to rule by decree and impose penalties for spreading false information [5, 37, 38].

The government's procurement of medical supplies was rife with anomalies, including purchasing a large number of ventilators from China (a key ally in Orbán's "Eastern opening" strategy) that were not fully utilized [10] and were still in storage in 2023 at prohibitive costs.[2] An EU-critical narrative also accompanied the vaccine policy. The European Commission planned a coordinated roll-out of vaccinations in all EU Member States on 27 December 2020. Still, Hungary started vaccinations a day early (on purpose), ignoring the plan [39].

The government had also authorized the use of Sinopharm and Sputnik V before they were approved by the European Medicines Agency [40, 41]. The lack of transparency and incomplete trial results raised concerns about the efficacy and safety of these vaccines. Despite these concerns, Hungary aimed to demonstrate its competence in handling the pandemic and use it to secure electoral support in the upcoming parliamentary elections [36].

The Eurosceptic stance of the government led Hungary to opt out of the EU vaccine deal with Pfizer and BioNTech, citing sufficient vaccine supply and high prices.

---

[2] https://telex.hu/belfold/2024/01/05/covid-lelegeztetogep-arveres-cegled-kulugy.

However, it was revealed that the Chinese vaccine was the most expensive of the available options.[3] Regardless, the Hungarian government criticized the EU's vaccine procurement process, claiming delays and failures while ignoring similar issues with shipments from Russia and China. In fact, by 2021, Hungary received more vaccines from Western manufacturers than from Eastern ones [42].

Even though the politicization of vaccine procurement posed risks for the government–as it could have undermined confidence in the vaccination process–geopolitics and vaccine diplomacy played a crucial role in the process [36, 40]. While the EU struggled to secure vaccine supplies, Russia and China used vaccine diplomacy to increase their influence in Central and Eastern Europe (CEE [41, 43]).

Hungary's close ties with Russia in the economic and political spheres, particularly in the energy sector, became increasingly evident [36]. The Orbán government rejected EU sanctions against Russia in connection with the invasion of Ukraine as it emphasized the importance of energy security [44]. Russian–Hungarian ties were especially strong in a regional comparison with Slovakia and the Czech Republic, let alone Poland [45]. For instance, in Slovakia, Prime Minister Igor Matovič resigned from his post due to the political crisis related to a secretly arranged procurement deal of Sputnik V vaccines as it was contradictory to EU rules and the stance of other government members [41].

Hungary also had close ties with China, as Chinese investment preferred Hungary over other Central and Eastern European countries [36]. Multiple decisions, such as agreements for the construction of the Budapest-Belgrade railway, a proposed campus for Fudan University [43, 46], and being the first signatory of China's Belt and Road Initiative in Europe [36], illustrate the favorable political climate and relations between China and Hungary.

Thus, it was unsurprising that Hungary was one of the primary targets of Chinese vaccine diplomacy in CEE [43]. The Orbán government was a key supporter of European cooperation with China, including Beijing policy initiatives such as mask diplomacy and vaccine diplomacy [46]. China's supply of vaccines to Europe was seen as a continuation of mask diplomacy and the Belt and Road Initiative, especially in the form of the 'Health Silk Road' [43].

The sentiment analysis of COVID-19-related content

Sentiment analysis refers to examining the opinions the reader forms of an object while interpreting texts [47]. During this identification process, the reader interprets words, phrases, expressions, text fragments, or extended text sections and classifies them into three collective sentiment types: positive, negative, and neutral [47]. Due to the proliferation of online news portals, creatin three collective sentiment types: positive, negative, and neutral [47]. Due to the proliferation of online news portals, creating ever-growing volumes of content, the automated sentiment analysis of political texts has become a widely used methodological choice in political communication and beyond [48–55].

---

3 https://qubit.hu/2021/05/19/az-astrazeneca-legolcsobb-a-sinopharm-a-legdragabb-a-vakcina-vilag
piacon.

As political discourse is seldom limited to a simple statement of information, the tone of a given text is at least as important as the facts. Moreover, tone can be the central component of individual decision-making and political judgment regarding the content of an article [55]. The sentiment analysis of COVID-19-related news coverage can shed light on how the public responds to the pandemic [56], especially as there is evidence of emotions being more important in the field of health care than in some other policy domains [57]. Sentiment analysis "has been widely used in public opinion monitoring" [58, 59], and automation is also helpful in teasing out subtle textual references at scale [60].

Jiang and Hyland [29] analyze the changing concerns of the international press and the central themes reflected in news media coverage. Through a linguistic analysis of a large corpus, the researchers identify key nouns, verbs, and phrases that highlight the evolving public interest in the pandemic. This study provides valuable insights into how the media has portrayed and reported on COVID-19 throughout 2020.

In one of the immediate precursors to our research design, Ghasiya and Okamura used a combination of topic modeling and sentiment analysis for COVID-19 news analysis [61]. In psychology, Aslam et al. [57] studied how COVID-19 impacted the mental health of the public and how such information was used to inform better policy-making. A study by Robertson et al. [62] demonstrated the opportunities to use automated multilingual text analysis to understand complex issues such as the pandemic. In one of the few available analyses related to the CEE region, researchers studied Croatian news with the help of topic modeling and a lexicon-based approach to detect sentiments [63]. Such studies, however, have not leveraged the supervised deep learning approach associated with large language models, which are considered to be the state-of-the-art approach to such problems.

As for our more immediate research topic, a branch of extant research focused specifically on sentiment related to vaccines. Ali et al. [64] investigated how the public viewed coronavirus vaccines in the US from early February to late March 2021 based on Twitter data. The analysis indicated that, despite the overall positive sentiment and increasing vaccination rates, negative sentiment towards COVID-19 vaccines persisted among segments of the population who were hesitant. The study underscored that specific population segments feared for their health in relation to COVID-19, which may have contributed to uncertainty related to taking up the vaccine. Based on these findings, the article highlighted the responsibility of governments to manage these fears through transparent information and communication."

Yin et al. [65] implemented a similar research design to ours with Twitter data collected from December 2020 to July 2021 related to seven popular vaccine brands (the five vaccines considered in this article plus Sinovac and Covaxin). Their results showed that most people had confidence in the effectiveness of the vaccines and were willing to be vaccinated. In contrast, negative tweets tended to be linked to news stories about post-vaccine deaths, vaccine shortages, and post-vaccine side effects.

Sear et al. [66] used machine learning to assess public pro- and anti-vax Facebook pages. They found that the debate in the anti-vaccine community was less focused on COVID-19 than in the pro-vaccine community. The anti-vax community, however, can

reach a larger group of individuals seeking online guidance on COVID-19 because it deals with a wider variety of topics.

Zhou et al. [67] examined social media discourse on COVID-19 vaccines by US politicians, health experts, and government agencies. They found that rising COVID-19 infection rates reduced the politicization tendencies of government agencies, and these agencies in "blue states" with a Democratic Party majority were more likely to politicize COVID-19 vaccines than those "red states" with Republican majorities. The authors considered blue state politicization to be an unexpected phenomenon and posited an incumbency effect as a possible root cause. In their understanding government communication managers consider public health figures as in-group allies in blue states while they were trying to disassociate themselves from politics in red states. The authors also cited the apparent lack of research on government communication, and as a general caveat, they also pointed to the fact that their methodology seemed to be more sensitive to the effect of politicization than other methods.

Béres et al. [68] also explored vaccine sentiment in their study based on Twitter data. Using a multilingual BERT model fine-tuned for sentiment analysis on product reviews in six languages (English, Dutch, German, French, Spanish, and Italian), they predicted the sentiment of tweets associated with the Pfizer-Biontech, Moderna, AstraZeneca, Sputnik V, and Sinopharm vaccines on a 1–5 star scale. They found the most positive sentiment for Moderna and the most negative for Sinopharm. They also showed a high geographic difference in opinions: European tweets were most negative towards Astra-Zeneca, while Asian ones were most critical towards Sinopharm.

Finally, Jalil et al. [19] studied the sentiment of tweets related to the COVID-19 pandemic using different feature sets and classifiers for sentiment classification to enable a better understanding and management of Twitter users' opinions. What is common in all these studies is that they mostly focused on Twitter for their data collection (which is understandable given the problems of securing Facebook data for research purposes). Yet Twitter is much less widely used in Hungary than in the United States and Western Europe.

As for our primary interest of Hungarian media, relevant studies are few and far between. In one such article, Szabó [69] analyzed the online news about the COVID-19 pandemic to examine whether the articles expressed emotive or factual content and focus. Results showed that nearly a quarter of online news content featured an emotional reaction to the pandemic or its consequences.

Hypotheses

Based on the overview of the politics and geopolitics of the Hungarian vaccination program, we formulate three hypotheses related to the treatment of vaccines in three media outlets—the first two concern saliency: the importance of vaccines on the media agenda. The third—a collection of three postulations—is related to sentiment driven by the domestic and geopolitical preferences of the outlets in question. We explore vaccine coverage in three outlets with a major share and/or ideological importance in the Hungarian online media market at the time:

- Origo[4]—a market-leading, right-wing site uncritical of the government,
- Alfahir[5]—a smaller, hard-right right-wing opposition site,
- Index/Telex[6] (simplified below as Telex)—a market-leading left-liberal opposition outlet.

In this period, media in Hungary had more limited freedom vis-á-vis liberal democracies, according to the 2021 World Press Freedom Index by Reporters Without Borders.[7]Therefore, selecting only the most popular or state-run media for analysis could lead to skewed results that rely too heavily on government-driven narratives. In creating our media sample, we took this factor into account. The collected corpus covers a broad spectrum of news portals. Origo and Index, besides tabloid outlets, were Hungary's two most prominent news portals at the time. Furthermore, the opposition was divided between left-liberal parties and a right-wing one (Jobbik). Therefore, we also included Alfahir, which, despite having fewer visits than the other two, was an important news source for Jobbik voters.

Our review of the geopolitics of vaccine roll-out showed that only Eastern vaccines had been widely available at the beginning of the period under study (March 2020–March 2022). Western vaccines were only offered on a massive scale from the second quarter of 2021, while the problem of vaccine procurement was constantly present in the political discourse. Therefore, in Hypothesis 1 (H1), we expect that the saliency of vaccines in the media is driven by the facts on the ground, especially regarding the intensity of the process of getting shots in the arms. We posit that the more people are directly affected by the vaccination program, the more media attention is dedicated to the problem.

> H1: In a monthly aggregation, there is a positive correlation between the count of first vaccine shots administered and the saliency of the vaccine topic for each media outlet.

For instance, we expect to find the highest number of vaccine mentions in the first quarter of 2021 in parallel with the launch and expansion of the vaccination program. With this hypothesis, we presume that vaccines appear in online media regardless of their provenance due to more active government communication on

---

[4] Origo.hu was founded in 1998 by the telecom company Matáv. It was later bought by New Wave Media in 2015 and then in 2018 by the media foundation KESMA (Central European Press and Media Foundation)—both closely linked to Orbán.

[5] Alfahir.hu is a right-wing opposition online portal owned by the Balanced Media Foundation, which is affiliated with the Jobbik Movement for Hungary, a right-wing populist party.

[6] Index.hu was founded in 1999 and has undergone several ownership changes. In 2020, Miklós Vaszily, with close ties to Viktor Orbán's friend and oligarch Lőrinc Mészáros, bought a 50% stake in the Indamedia group, which owns index.hu. Following the resignation of the editor-in-chief, most of the portal's journalists left index.hu. In the second half of 2020, they founded a new media outlet, which eventually became telex.hu. In the Telex dataset, we've included Index articles between March and August 2020, as these articles were written by the founding staff of Telex. There are no articles for September 2020 for this outlet, as Telex only launched a month later.

[7] https://rsf.org/en/index?year=2021.

the subject (as it may consider such communication an important tool for disease control).

Our second hypothesis concerns the relative importance of the vaccine topic by outlet. Further research suggests that media coverage of policy issues follows a punctuated equilibrium pattern [23]. This indicates that external factors or policy dynamics prompt shifts in media focus and policy agendas. Applying this framework to the case at hand means that we expect that all media outlets, regardless of ideological leaning, will cover the pandemic and vaccination as a key topic in this time frame.

H2: The Saliency of the vaccine topic does not show a statistically significant difference between the outlets.

Finally, even as we presume saliency to show similar characteristics across outlets, we posit that coverage will diverge in terms of sentiment in the news portals in the sample. As Telex was a Western-oriented, government-critical outlet in the period, we assume it to be the most critical of the government's purchases of vaccines from Russia and China. Origo was a government-friendly outlet. Therefore, we assume that it aimed to legitimize the government's vaccine strategy and thus to be supportive of both Eastern vaccines. Alfahir, as a right-wing opposition portal, is expected to be more critical of Chinese and Western vaccines because of the close Russian connections of Jobbik, its parent organization in party politics, at the time.

Therefore, given these geopolitical preferences of the outlets in question, with the three parts of Hypothesis 3, we explore the sentiment of coverage in each sample outlet for each vaccine. All three posit a rank order in the net positivity (see below) of the coverage of each of the three vaccines/vaccine groups in each outlet.

- H3a: The rank order of the net positivity coverage for Western vaccines is Telex > Alfahir > Origo.
- H3b: The rank order of the net positivity coverage for the Russian vaccine is Alfahir > Origo > Telex.
- H3c: The rank order of positive coverage for the Chinese vaccine is Origo > Telex > Alfahir.

H1 is operationalized by investigating the monthly cross-correlations between the number of first vaccine shots administered and the number of sentences mentioning a vaccine for each outlet-vaccine shots pairing. As we look for the higher coefficient, we look at three lags in our regression analysis: zero (synchronicity), plus one (sentences from month t are correlated to vaccination levels to month $t-1$), and minus one (vice versa).

We operationalize H2 by calculating

1. the global mean of the monthly rate of the saliency of the vaccine topic for each of the three outlets is defined as

$$\frac{Number\ of\ sentences\ mentioned\ vaccines}{Number\ of\ sentences\ in\ articles\ mentioning\ COVID\text{-}19},$$

2. compare saliency rate AR(1) autoregressive models between outlets.

We only consider COVID-related news items (having at least a single mention of the term "COVID" or "coronavirus") to craft a text corpus with less noise and normalize vaccine-related news within all pandemic-relatede types of vaccines (Western, Russian, Chinese).

We adopt a conventional codebook developed by Plutchik [15], which aggregates his original eight emotions (not containing "neutral") into two sentiment categories: positive ("Joy", "Surprise", "Trust"), negative ("Anger", "Disgust", "Fear", "Sadness"), and we added neutral category too (the Plutchik's "Anticipation" category, was not included in our corpus). We calculated net positivity by subtracting negatives from positives and dividing this score by the total (positives, negatives, neutrals). Thus, net positivity shows an overrepresentation of positivity expressed over the total base.

Data and methods

We examined three corpora covering major news sites in Hungary between March 2020 and March 2022: Origo, Telex, and Alfahir. Data collection was focused on online news site article texts from March 1st, 2020, to March 31st, 2022 (except Telex for September 2020). These starting and ending months marked major milestones in the Hungarian COVID-19 policy regime: March 2020 saw the first wave of the pandemic and the accompanying curfew, while in March 2022, most major restrictions were rolled back, and the count of first shots administered plummeted.

Table 1 shows that sentences were unevenly distributed between the three outlets, and the prevalence of sentences mentioning vaccines was also high. They often referred to a single brand, while multiple references were common. In this article, we only focus on single-mention sentences as these comprise the bulk of the database and are easier to handle from a research design perspective.

Sentences mentioning a single vaccine brand (and a single vaccine brand only) were extracted from article texts and fed into a BERT-based classification model fine-tuned for our project. BERT is short for Bidirectional Encoder Representations from Transformers, a widely used language model for many applications, such as classification and next-sentence prediction [70]. BERT models are pre-trained on large corpora to understand backward and forward contextual information from texts.

BERT models are extensive but must be fine-tuned on specific tasks to be precise and robust for the given application [71, 72]. For this paper, we used a new model developed by the co-authors called emBERT [72]. This model uses HUBERT as a base, trained on a large-scale Hungarian corpus [73]. EmBERT, in turn, was fine-tuned for a manually coded sentence-level media dataset dating from before the

Table 1 Articles and sentences mentioning vaccines (by outlet)

Outlet	Articles collected	Count of COVID-19 articles	Sentences in COVID-19 articles	Sentences of single mentions	Sentences of multiple mentions
Origo	142,423	27,318	458,710	4712	487
Alfahir	24,404	7120	100,016	1594	135
Telex	66,121	19,688	594,749	6832	550
TOTAL	232,948	54,126	1,153,475	13,138	1172

pandemic. The model can classify sentences into seven categories (Happiness, Success, Trust, Anger, Disgust, Fear and Neutral).

The emBERT model is unique in size, quality, and granularity in the emotion/ sentiment analysis domain. The fine-tuning dataset comprised 14,637 sentence-length news article excerpts from the Hungarian news site Index, dating from 2000 to 2016. The dataset was manually annotated using the adapted Plutchik codebook.

The original fine-tuning of the eight emotions resulted in a weighted average accuracy and F1-score[8] of 0.74 and 0.73, respectively [74]. This score is close to the 0.75 F1 threshold, which is often considered to be a "research-grade" output [74].

With the fine-tuned emBERT model, we classified emotions in the sentences, which were then aggregated into positive (Happiness, Success, Trust), negative (Anger, Disgust, Fear, Sadness), and neutral (Neutral) sentiments. Over the general sentiment analysis, emotion analysis is a more complex way of classifying opinions as we move over the general distribution by studying the specific emotions of the texts. There are examples of more detailed emotion analysis combined with less detailed sentiments in the literature. The emotion categories of happiness, joy, and trust can be aggregated into emotion categories as positive emotions, while fear, anger, and disgust can be aggregated into emotion categories as negative emotions [75, 76]. As a validity check, manual double-blind validation of aggregated sentiment classes was done on a simple random sample of 10% of the single-mention rows of data.

This validation process yielded a weighted accuracy of 0.85, with an inter-annotator agreement (Cohen's kappa) of 0.8, which signals almost perfect agreement (see [77, p. 362], on the definition of "substantial agreement"). As the BERT classification was not strictly target-dependent (there was no rule that the sentiment references of the sentence should be targeted on the vaccine brands themselves), validation also included a decision on whether the sentiment target was a vaccine. This additional validation showed a target accuracy of 80%.

After separating single- and multiple-mention sentences, we first validated the single/multi-mention separation on a 20% sample with expert validation, which showed an accuracy of 88%. We excluded short, unintelligible phrases based on

---

[8] The weighted macro F1 score is a widely used metric for evaluating model performances in classification tasks. The higher the value, the better the performance. There is no definite standardization on what is counted as state-of-the-art; cutoff points are determined based on the task, language etc. in question [74].

their character length as they turned out to be the result of improper character coding. As a final validation step, we have validated vaccine mentions based on sentences actually mentioning vaccines. This reached a 100% accuracy on a random sample of 100 observations.

We also used search functions to filter out texts where Sputniks referred to the 1959 space probe or the news agency. Table 2 shows translated sample sentences for each sentiment category from the pre-processed corpus.

Descriptive statistics

The descriptive statistics of the vaccine coverage of the three outlets in the sample show significant variation. As Table 3 presents, large deviations were observable for the count of COVID- and vaccine-related sentences but not the average length of COVID-related articles measured in words. While Origo averaged 39% more COVID-19-related articles per month than Telex, the latter still published 30% more vaccine-related sentences. In the case of Alfahir, there is an evident concentration of more vaccine coverage in fewer general COVID sentences. These data are a clear sign of an editorial focus on vaccines in the case of Alfahir.

Additionally, the data generation process was intertwined with the history of the pandemic: as cases piled up, coverage also increased. This is visible in Figs. 1, 2, and 3. These show that most mentions for any of the three vaccine types gathered steam in late 2020, albeit on different levels (most vaccines were concurrently unveiled in 2020). Western vaccine mentions are the most numerous around March 2021, and they had (a significantly smaller) second peak in late 2021.

Such a second wave is barely discernible for the other two vaccines. The figures also reflect that around the Summer of 2021, the Western vaccine supply became abundant, and anyone who had not been vaccinated could receive it. By this time, the vaccination program was over its zenith: the first 6 months of 2021

Table 2 Sample sentences for each sentiment per outlet

Sentiment	Outlet	Example
Positive	Alfahir	Moderna was certified by health agencies in several countries
	Telex	Moderna announced a safe vaccine
	Origo	The Chinese Sinopharm vaccine was tested in Hungary and proved harmless but effective
Negative	Alfahir	Earlier, a shipment of AstraZeneca was halted due to the death of a soldier in Sicily
	Telex	The latest debate is whether the Chinese vaccine is better against new variants
	Origo	US-based Moderna can also only send a limited amount
Neutral	Alfahir	As is known, the vaccine from Pfizer and Moderna was also prepared using this method
	Telex	BioNTech and Pfizer's vaccine technology is based on messenger RNA (mRNA) molecules
	Origo	The decision covers the use of Pfizer and Moderna vaccines

Table 3 The saliency of the vaccine topic (per month values)

Outlet	The mean number of COVID-19 related sentences	The mean number of words in COVID-19 related articles	Mean no. of vaccine sentences
Alfahir	4001	21	72
Origo	18,348	23	208
Telex	24,781	21	308

saw a monthly average of 913,000 first shots, while the last 6 months did not reach 130,000 per month (see also Fig. 4).

As for sentiment, as Table 4 shows, the overall sentiment category is shared in the outlets for sentences with a mention of a vaccine. Ratios are similar for Alfahir and Origo, while Telex was less positive in its vaccine coverage.
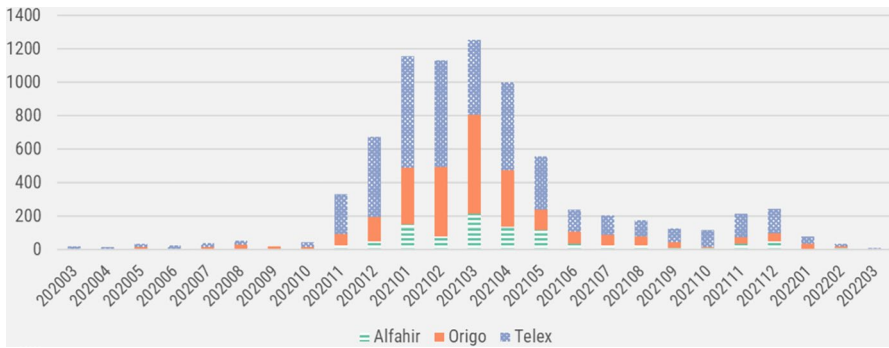
Fig. 1 Counts of Western vaccine-related sentences per outlet
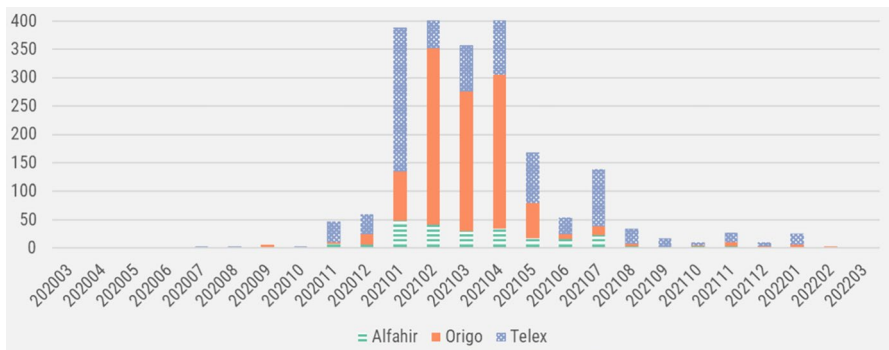
Fig. 2 Counts of Chinese vaccine-related sentences per outlet
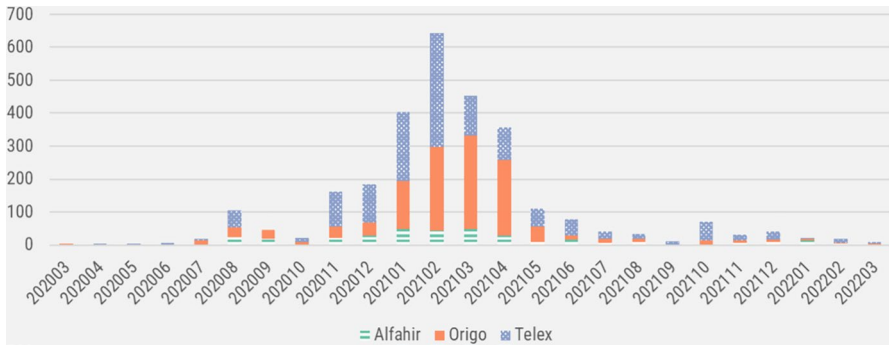
Fig. 3 Counts of Russian vaccine-related sentences per outlet

Results

The empirical research design was aimed at investigating three hypotheses. H1 posited a positive correlation between the count of first vaccine shots and the saliency of the vaccine topic. Figure 4 presents general trends that already hint at their close interrelation. The left Y-axis shows the monthly rate of vaccine mentions from the respective outlet's COVID-related articles. The right Y-axis shows the monthly count of first vaccine shots. We calculated cross-correlations for the first differences (due to serial correlation in the time series) to examine this thesis at −1, 0, and 1 lags (see Table 5). Short lags are a staple of agenda and public opinion research, as exhibited in several comprehensive studies, such as
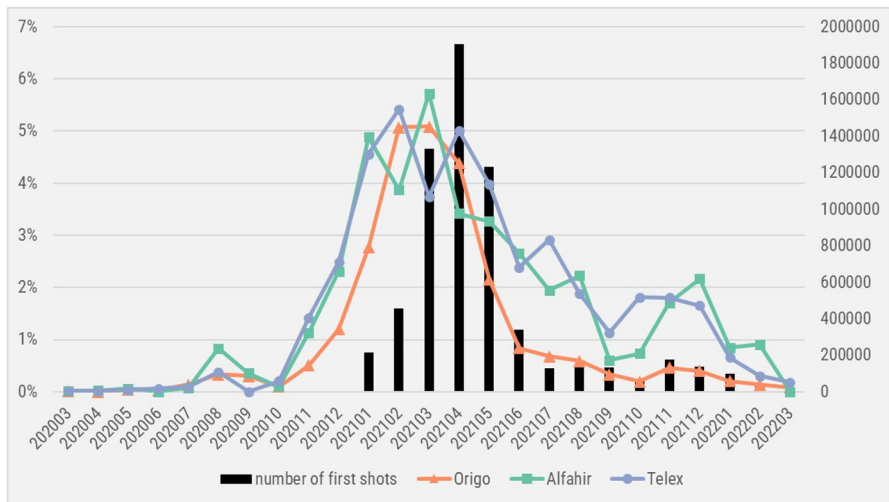
Fig. 4 Vaccine-saliency by outlets and vaccination trends (Number of first shots on right, vaccine-related sentence rate on left. COVID-related sentences are defined as sentences from articles containing references to COVID. This provides a normalized denominator)

Table 4 Sentiment shares of sentences in outlets (absolute counts in brackets)

	Positive share	Negative share	Neutral share
Alfahir	65% (1031)	27% (425)	9% (138)
Origo	66% (3106)	25% (1159)	9% (447)
Telex	58% (3971)	29% (1976)	13% (885)

Vliegenthart et al. [78], Froio et al. [79], and Lovett et al. [80]. The lags in this case shed light on whether the media coverage precedes vaccination dynamics (-1), coincides with them (0), or trails vaccination (1). In this case, the lags are useful for uncovering possible shifts caused by intrinsic agenda dynamics or data grouping.

Data shows the offset of the vaccine coverage can be significantly measured for Origo and Alfahir at -1 and 0 lags. The attention dedicated to the topic in Origo and Alfahir precedes or is in step with the first vaccine shots. The coefficient size favors -1 for Origo, and 0 for Alfahir.

Both outlets' coefficients are significant even after adjustment for multiple comparisons at the 0.05 level and, therefore, partially support H1. General trends are shared between the outlets, as all exhibit (insignificant) cross-correlation coefficients ranging between 0.1 and 0.438 for lags 0 and -1.

For testing our second hypothesis, we compared the saliency rates of outlets to see if they show statistically significant differences. Despite differences between certain aspects of the outlets, we hypothesized no significant differences in saliency dynamics overall. Figure 4 shows Origo's markedly different but ultimately insignificant path compared to the other two sites. Vaccine mentions in all three outlets peaked during the spring of 2021, coinciding with the bulk of the first vaccine shots in the population. However, vaccine saliency in Origo peaked shorter and lower. (After that, it followed the trend in vaccine administration to the population in a steep fall, just like the other two outlets.)

The news portals critical of the government kept the 'buzz' (respective to their overall article volumes) running for longer before reaching more or less the same endpoint. These visual observations hold even in the aggregate, as illustrated by Table 6.

Data shows that the saliency rate is highest for Telex, with Alfahir coming a close second. Pairwise F-tests between AR(1) models of the datasets show that no pairs show statistically significant differences between their dynamics (see Tables 6 and

Table 5 Cross-correlation between first vaccine shots and vaccine-related sentences

Outlet	-1 lag (coverage preceding)	0 lag (no offset)	1 lag (coverage trailing)
Origo	0.917 (0.000007)	0.474 (0.04)	-0.262 (0.51)
Alfahir	0.408 (0.02)	0.438 (0.03)	-0.062 (0.61)
Telex	0.130 (0.2)	0.100 (0.75)	-0.004 (0.96)

Adjusted p-values in brackets. Data provided at various lags. Familywise error rate was controlled by Holm-Bonferroni adjustments throughout the paper at the 0.05 level

Table 6 Means of monthly rates of vaccine saliency and AR(1) model parameters

Outlet	Mean with SD	Intercept (SE)	AR1 (SE)
Origo	1.04% (1.6%)	0.0015 (0.002)	0.8638 (0.102)
Alfahir	1.6% (1.6%)	0.0037 (0.003)	0.7789 (0.128)
Telex	1.69% (1.7%)	0.0026 (0.003)	0.8551 (0.104)

Table 7 Results of pairwise F-tests on AR(1) models

Comparison	Pairwise F-tests on AR(1) models	
	F (df1, df2)	adj. *p*-value
Origo-Alfahir	0.617 (24, 24)	0.88
Origo-Telex	0.826 (24, 24)	0.68
Alfahir-Telex	0.747 (24, 24)	0.76

7). Even though the autoregressive model shows no significant differences between the internal dynamics of the outlets, Origo seems to precede the vaccination dynamics by 1 month. This difference can be explained by the steeper fall of this outlet after the peak, as this is the only obvious difference between the outlets' dynamics. This partially refutes our second hypothesis.

The three parts of H3 posited a geopolitical explanation for the differences in vaccine coverage between the three outlets in the sample. We presumed that Telex would carry a mostly positive tone towards Western vaccines. At the same time, a preference was expected for the Russian vaccine in Alfahir and the Chinese one in Origo. Figure 5 displays marked differences in the sentiments of the coverage of various vaccines.

The Chinese Sinopharm is the least favored vaccine in all three outlets. Alfahir was particularly critical, with a net positivity rate below 10% (meaning that negative sentences were close to outweighing positives). Alfahir also stands out with its preference for Russian vaccines, even in comparison to Western ones (otherwise leading the pack).

An even more telling visualization can focus on relative net positivity (see Fig. 6): a measure of outlet-specific net positivity of a vaccine compared to generic (ungrouped) net positivity. This allows us to control for differences in general positivity levels (since e.g., Origo is 41% more positive than Telex, making it by far the most positive of the three outlets).[9] These relative relationships are mostly unchanged compared to the absolutes, but the values provide a more accurate representation of the logic of the data.

Our results related to the three parts of H3 are presented in Table 8. Hypothesis 3a was refuted, as Origo carried a more positive coverage of Western vaccines vis-á-vis the other two. We supported Hypothesis 3b, as Alfahir was by far the most positive with the Russian vaccine, followed by Origo. The 19% relative net positivity

---

[9] A case could be made for using absolute values for a study that focuses on agenda effects on the general public in which one should also control for the reach (readership) of the outlets.
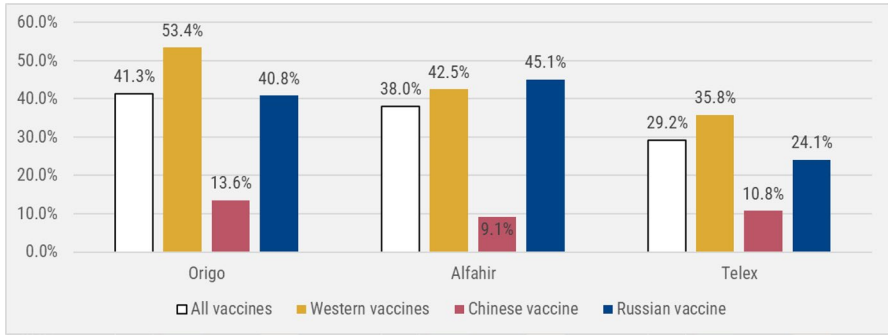
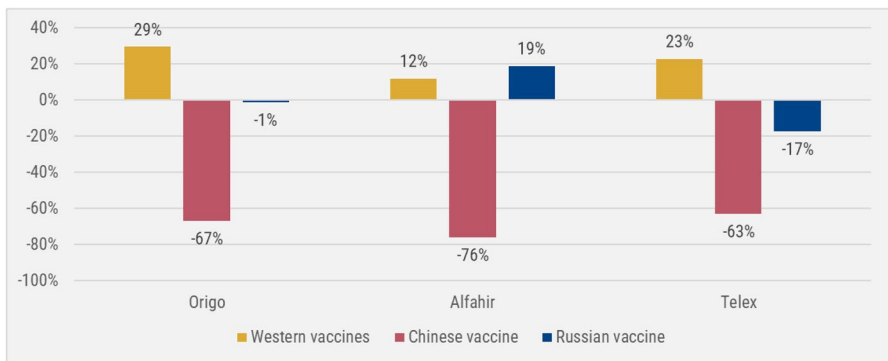**Fig. 5** Net positivity by outlets and vaccine types

Fig. 6 Relative net positivity of outlets toward vaccines

Table 8 Results for relative net positivity (H3)

Vaccine type	Predicted order	Empirical order	Hypothesis
Western (H3a)	Telex > Alfahir > Origo	Origo > Telex > Alfahir	Refuted
Russian (H3b)	Alfahir > Origo > Telex	Alfahir > Origo > Telex	Supported
Chinese (H3c)	Origo > Telex > Alfahir	Telex > Origo > Alfahir	Refuted

rate in Alfahir is a clear indicator of a strong editorial focus on making Sputnik V look good.

Finally, our analysis also refuted H3c. Even though Alfahir produced articles with drastically negative sentiments towards the Chinese vaccine compared to the other outlets, Telex came to the forefront as being less negative than the other, even though not by gigantic margins. This extremely low rate from Alfahir (absolute score: below 10%, relative score: −76%) was the worst in our dataset. Origo also produced less positive coverage of the Chinese vaccine than Telex, which rounds up the analysis of H3c.

Discussion

In this article, we analyzed the coverage of various pro-government and Orbán-critical media vaccine brands to examine the differences in their treatment of Eastern and Western vaccines. We trained a large language model (emBERT). We relied on state-of-the-art machine learning analysis to investigate all COVID-19-related news (72 339 articles) published on three major Hungarian news portals between March 2020 and March 2022.

Our results show a positive correlation between the number of sentences for two out of three outlets mentioning one of a vaccine and the (first) shots administered and a non-significant difference between the vaccine attention rates of the analyzed outlets. The pro-government portal showed more positivity towards Western vaccines than a hard-right, anti-government one. This latter, Alfahir, also produced more positive reports about the Russian vaccine, while its coverage was more negative towards the Chinese vaccine [5, 11, 14]. These findings align with existing studies that discuss the influence of media on public health narratives and suggest that media coverage can precede and potentially steer public health interventions like vaccination programs [26, 29].

We also supported a theme in the literature about the precedence of media coverage over-vaccination program roll-outs. Not all outlets treat vaccines as a premier topic in their agenda. The government-affiliated Origo lagged behind both in terms of timing and level of attention. Moreover, while the pro-government media exhibited support for Eastern vaccines, their coverage was surprisingly positive towards Western vaccines as well. This pattern likely reflects a strategic initiative to maintain public confidence in the government's pandemic response, ensuring broad acceptance of available vaccines, regardless of their origin [36, 40]. In contrast, the distinct bias of Alfahir towards Russian vaccines can be attributed to its political ties, underscoring the influence of geopolitical affiliations on media behavior [41, 45].

This comprehensive media analysis sheds light on how news outlets have navigated the complex landscape of vaccine diplomacy and public health, influenced by domestic and international political dynamics [43, 46]. By applying advanced text-mining techniques and sentiment analysis, we have unveiled nuanced patterns of media influence that dictate public perception and acceptance of health policies, suggesting a profound interconnection between media portrayal and geopolitical stances.

A straightforward interpretation of these results could point to an intent on behalf of the government to tamp down the risks of the pandemic in the public eye and divert attention away from potential hiccups and controversies with vaccine procurement (of which there was plenty when it came to participation in an EU-program aimed at securing Western vaccines).

At the same time, opposition media was particularly active in the buildup phase of the pandemic and focused on the topic more intensively and longer after the vaccination program peaked. Overall, coverage in the diverse Hungarian media landscape was closely interlinked with real-life processes and betrayed a generally positive outlook towards vaccines (except the Chinese vaccine in Alfahir). This state of

affairs showed a contrast with more vaccine-critical treatments in the media of some other developed countries [26–28, 30].

Regarding the geopolitical dimension of the discourse around vaccines, our first finding was that the pro-government portal, while supporting and promoting the Eastern vaccines, was not at all opposed to the Western ones. This meant that their overall appearance was positive beyond our expectations. This finding is assumed to be due to the government's aim of keeping the epidemic under tight control, in which the success of the vaccination program was a key element.

The second result concerned the anti-government, right-wing portal (Alfahir), which was the most positive towards the Russian vaccine and much more negative towards the Chinese vaccine. These results can be explained by the Russian political connection of the party behind the portal. This is likely because the government was focused on maintaining strict control over the epidemic, and the success of the vaccination program was a crucial factor in achieving this goal.

Our analysis has shown the possibilities of statistical, text-mining analysis of media text to reveal hidden patterns. While also offering a dataset, a fine-tuned model and an analytical framework for future research. In addition, by applying sentiment analysis to vaccine-related news, we could indirectly reveal the geopolitical preferences of the parties behind the news portals. The research design can be further clarified by hand-coding sentences in which multiple mentions were made, which would allow for further cross-checking of how vaccine groups were covered.

Data availability The datasets analyzed during the current research are available in the Open Science Framework repository at https://osf.io/h4b7a/.

Declarations

Conflict of interest The author(s) declared no potential conflict of interest.

References

1. Bernek, Á. (2018). Hazánk keleti nyitás politikája és a 21. Századi geopolitikai stratégiák összefüggései. *Külügyi Szemle, 17*(2), 122–144.
2. Krekó, P. (2018). Oroszország a magyar közvéleményben. In T. Kolosi & I. G. Tóth (Eds.), *Társadalmi Riport 2018* (pp. 382–396). TÁRKI Social Research Institute, Budapest. https://www.tarki.hu/sites/default/files/trip2018/382-396_Kreko_Oroszorszag_a_magyar_kozvelemenyben.pdf
3. Bajomi-Lázár, P., & Stępińska, A. (2019). The Hungarian media system under political pressure. *Central European Journal of Communication, 12*(2(23)), 257–262. https://doi.org/10.19195/1899-5101.12.2(23).9
4. Polyák, G. (2019). Media in Hungary: Three pillars of an illiberal democracy. In E. Połońska & C. Beckett (Eds.), *Public service broadcasting and media systems in troubled European democracies* (pp. 279–303). Springer. https://doi.org/10.1007/978-3-030-02710-0_13
5. Krekó, P. (2020, April 24). *Orbán's pandemic authoritarian grab | Heinrich Böll Stiftung | Prague Office—Czech Republic, Slovakia, Hungary*. Heinrich-Böll-Stiftung. https://cz.boell.org/en/2020/04/24/orbans-pandemic-authoritarian-grab
6. Polyák, G., & Nagy, K. (2021). Regulating health communication in the post-truth era. *Intersections, 7*(3), 120–138. https://doi.org/10.17356/ieejsp.v7i3.813
7. Lamour, C. (2023). Orbán placed in Europe: Ukraine, Russia and the radical-right populist heartland. *Geopolitics*. https://doi.org/10.1080/14650045.2023.2241825
8. Bíró-Nagy, A., & Szászi, Á. J. (2023). The roots of COVID-19 vaccine hesitancy: Evidence from Hungary. *Journal of Behavioral Medicine, 46*(1–2), 185–200. https://doi.org/10.1007/s10865-022-00314-5
9. Marton, P., Matura, T., & Somogyvári, C. (2023). "Dracunculus against the dragon": Hungarian Prime Minister Viktor Orbán's public vaccination as simultaneous enactment of public health and foreign policy. *Journal of Contemporary Central and Eastern Europe, 31*(2), 409–428. https://doi.org/10.1080/25739638.2023.2221923
10. Szopkó, Z. (2021, November 18). Hungary donated coronavirus vaccines and ventilators worth HUF 19 billion abroad. *Átlátszó.Hu English*. https://english.atlatszo.hu/2021/11/18/hungary-donated-coronavirus-vaccines-and-ventilators-worth-huf-19-billion-abroad/
11. Griffen, S. (2020). Hungary: A lesson in media control. *British Journalism Review, 31*(1), 57–62. https://doi.org/10.1177/0956474820910071
12. Anderson, N., Commins, A., & Whitten-Woodring, J. (2023). Fostering bottom-up censorship from the top-down: Nationalism and media restrictions. *The International Journal of Press/Politics*. https://doi.org/10.1177/19401612231172306
13. Bellucci, L. (2021). Media law, illiberal democracy and the COVID-19 pandemic: The case of Hungary. In M. Deflem & D. M. D. Silva (Eds.), *Sociology of crime, law and deviance* (pp. 151–167). London: Emerald Publishing Limited. https://doi.org/10.1108/S1521-613620210000026010
14. Bleyer-Simon, K. (2021). Government repression disguised as anti-disinformation action: Digital journalists' perception of COVID-19 policies in Hungary. *Journal of Digital Media and Policy, 12*(1), 159–176. https://doi.org/10.1386/jdmp_00053_1
15. Plutchik, R., & Kellerman, H. (Eds.). (1980). *Emotion: Theory, research, and experience* (Vol. 1). Academic Press.
16. Mayer, M., Bichler, B. F., Pikkemaat, B., & Peters, M. (2021). Media discourses about a super-spreader destination: How mismanagement of Covid-19 triggers debates about sustainability and geopolitics. *Annals of Tourism Research, 91*, 1–15. https://doi.org/10.1016/j.annals.2021.103278
17. Zunino, E. (2021). Digital Media and COVID-19: Overinformation, polarization and misinformation. *Universitas-XXI, Revista de Ciencias Sociales y Humanas, 34*, 133–154.
18. Sidarta, C., Kurniawan, A., Lugito, N. P. H., Siregar, J. I., Sungono, V., Heriyanto, R. S., Sieto, N. L., Halim, D. A., Jodhinata, C., Rizki, S. A., Sinaga, T. D., Marcella, C., Wijovi, F., Susanto, B., Marcella, E., Wibowo, J., Anurantha, J. J., Imanuelly, M., Indrawan, M., & Heryadi, N. K. (2022). The determinants of COVID-19 vaccine acceptance in sumatra. *Kesmas: Jurnal Kesehatan Masyarakat Nasional, 17*(1), 32–39. https://doi.org/10.21109/kesmas.v17i1.4958
19. Jalil, Z., Abbasi, A., Javed, A. R., Badruddin Khan, M., Abul Hasanat, M. H., Malik, K. M., & Saudagar, A. K. J. (2022). COVID-19 related sentiment analysis using state-of-the-art machine learning and deep learning techniques. *Frontiers in Public Health, 9*, 1–14. https://doi.org/10.3389/fpubh.2021.812735

20. Ruf, T., Song, J., & Zhang, B. (2021). Political relations and media coverage. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3849962

21. Golan, G. J., Waddell, T. F., & Barnidge, M. (2021). Competing identity cues in the hostile media phenomenon: Source, nationalism, and perceived bias in news coverage of foreign affairs. *Mass Communication and Society, 24*(5), 676–700. https://doi.org/10.1080/15205436.2021.1884263

22. Karameti, A., & Bilalli-Zendeli, A. (2020). News media and political agenda during the covid 19 pandemic. *Justicia—International Journal of Legal Sciences, 13–14*, 23–28.

23. Baumgartner, F. R., Jones, B. D., & Mortensen, P. B. (2018). Punctuated equilibrium theory: Explaining stability and change in public policymaking. In C. M. Weible & P. A. Sabatier (Eds.), *Theories of the policy process* (4th ed., pp. 55–101). London: Routledge. https://doi.org/10.4324/9780429494284

24. Boda, Z., & Patkós, V. (2018). Driven by politics: Agenda setting and policy-making in Hungary 2010–2014. *Policy Studies, 39*(4), 402–421. https://doi.org/10.1080/01442872.2018.1478075

25. Sebők, M., & Boda, Z. (Eds.). (2021). *Policy agendas in autocracy, and hybrid regimes: The case of Hungary*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-73223-3

26. Faccin, M., Gargiulo, F., Atlani-Duault, L., & Ward, J. K. (2022). Assessing the influence of French vaccine critics during the two first years of the COVID-19 pandemic. *PLoS ONE, 17*(8), 1–19. https://doi.org/10.1371/journal.pone.0271157

27. Hüsemann, F. (2021). *A cure that no one wants? A comparative content analysis on how different German and UK newspapers frame the Covid-19 vaccine AZD1222 by Astrazeneca over time* (public). http://essay.utwente.nl/86507/

28. Jamil, S., & Appiah-Adjei, G. (2020). Battling with infodemic and disinfodemic: The quandary of journalists to report on COVID-19 pandemic in Pakistan. *Media Asia, 47*(3–4), 88–109. https://doi.org/10.1080/01296612.2020.1853393

29. Jiang, F. K., & Hyland, K. (2022). COVID-19 in the news: The first 12 months. *International Journal of Applied Linguistics, 32*(2), 241–258. https://doi.org/10.1111/ijal.12412

30. Motta, M., & Stecula, D. (2023). The effects of partisan media in the face of global pandemic: How news shaped COVID-19 vaccine hesitancy. *Political Communication, 40*(5), 505–526. https://doi.org/10.1080/10584609.2023.2187496

31. Akcakir, G., Jiang, Y., Luo, J., & Noh, S. (2023). Validating a mixed-method approach for multilingual news framing analysis: A case study of COVID-19. *Computational Communication Research, 5*(2), 1–70. https://doi.org/10.5117/CCR2023.2.11.AKCA

32. Bailard, C. S. (2023). Do local newspapers mitigate the effects of the polarized national rhetoric on coviD-19? *The International Journal of Press/Politics, 28*(4), 929–951. https://doi.org/10.1177/19401612211072774

33. De León, E., Makhortykh, M., Gil-Lopez, T., Urman, A., & Adam, S. (2023). News, threats, and trust: How COVID-19 news shaped political trust, and how threat perceptions conditioned this relationship. *The International Journal of Press/Politics, 28*(4), 952–974. https://doi.org/10.1177/19401612221087179

34. Papp, Zs., & Nkansah, G. B. (2023). The political component of COVID-19 vaccine choice: Results from a conjoint experiment. *Public Health, 217*, 33–40. https://doi.org/10.1016/j.puhe.2023.01.014

35. Yu, H., Lu, H., & Hu, J. (2021). A corpus-based critical discourse analysis of news reports on the COVID-19 pandemic in China and the UK. *International Journal of English Linguistics, 11*(2), 36. https://doi.org/10.5539/ijel.v11n2p36

36. Peragovics, T., & Kállai, P. (2021). Hungary: No need to preach to the choir. In *Responding to the COVID-19 'infodemic'. The Hague Program for Cyber Norms, The Hague* (pp. 33–41). The Hague Program for Cyber Norms. http://real.mtak.hu/134268/

37. Guasti, P., & Bustikova, L. (2022). Pandemic power grab. *East European Politics, 38*(4), 529–550. https://doi.org/10.1080/21599165.2022.2122049

38. Kovács, K. (2021, March 11). Hungary and the pandemic: A pretext for expanding power. *Verfassungsblog*. https://verfassungsblog.de/hungary-and-the-pandemic-a-pretext-for-expanding-power/

39. Hernandez, A. (2020, December 26). Hungary starts vaccinating one day early, against EU plan. *POLITICO*. https://www.politico.eu/article/hungary-administers-covid-19-vaccine-to-health-workers/

40. Frenyó, A. (2021, February 20). *Hungary: Pro-government scare tactics to intimidate critical voices amid vaccination rollout | Heinrich Böll Stiftung | Prague Office—Czech Republic, Slovakia, Hungary*. Heinrich-Böll-Stiftung. https://cz.boell.org/en/2021/02/20/hungary-pro-government-scare-tactics-intimidate-critical-voices-amid-vaccination-rollout

41. Gosling, T. (2021, March 31). Russia and China Are Exploiting Europe's Vaccine Shortfalls. *Foreign Policy*. https://foreignpolicy.com/2021/03/31/russia-china-vaccine-diplomacy-slovakia-europe-eu-slow-rollout/

42. Zsiros, S. (2021, November 26). *Hungary silently rejoins the EU's joint vaccine pool*. Euronews. https://www.euronews.com/my-europe/2021/11/26/hungary-silently-rejoins-the-e-u-s-joint-vaccine-pool

43. Atamanenko, A., & Konopka, N. (2022). Russia's and China's vaccine diplomacy in central and eastern Europe. *Copernicus Political and Legal Studies, 1*(1), 7–13. https://doi.org/10.15804/CPLS.20221.01

44. Dalos, G. (2022, October 29). Orbán says Hungary is 'exempt' from the conflict: Tell that to his friend in Moscow. *The Guardian*. https://www.theguardian.com/commentisfree/2022/oct/29/viktor-orban-hungary-conflict-moscow-war-in-ukraine

45. Guasti, P., & Bílek, J. (2022). The demand side of vaccine politics and pandemic illiberalism. *East European Politics, 38*(4), 594–616. https://doi.org/10.1080/21599165.2022.2122047

46. Paszak, P. (2021, April 5). China's vaccine and mask diplomacy in Hungary. *Warsaw Institute*. https://warsawinstitute.org/chinas-vaccine-mask-diplomacy-hungary/

47. Liu, B. (2010). Sentiment analysis and subjectivity. In N. Indurkhya & F. J. Damerau (Eds.), *Handbook of natural language processing* (2nd ed., pp. 627–666). London: Chapman and Hall/CRC. https://doi.org/10.1201/9781420085938

48. Boukes, M., Van De Velde, B., Araujo, T., & Vliegenthart, R. (2020). What's the tone? Easy doesn't do it: Analyzing performance and agreement between off-the-shelf sentiment analysis tools. *Communication Methods and Measures, 14*(2), 83–104. https://doi.org/10.1080/19312458.2019.1671966

49. Haselmayer, M., & Jenny, M. (2017). Sentiment analysis of political communication: Combining a dictionary approach with crowdcoding. *Quality and Quantity, 51*(6), 2623–2646. https://doi.org/10.1007/s11135-016-0412-4

50. Mohammad, S. M. (2016). Sentiment analysis: Detecting valence, emotions, and other affectual states from text. In *Emotion measurement* (pp. 201–237). Elsevier. https://doi.org/10.1016/B978-0-08-100508-8.00009-6

51. Mullen, T., & Malouf, R. (2006). A preliminary investigation into sentiment analysis of informal political discourse. *Computational Approaches to Analyzing Weblogs*, 159–162. http://www.aaai.org/Library/Symposia/Spring/2006/ss06-03-031.php

52. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135. https://doi.org/10.1561/1500000011

53. Rauh, C. (2018). Validating a sentiment dictionary for German political language—A workbench note. *Journal of Information Technology and Politics, 15*(4), 319–343. https://doi.org/10.1080/19331681.2018.1485608

54. Van Atteveldt, W., Kleinnijenhuis, J., Ruigrok, N., & Schlobach, S. (2008). Good news or bad news? Conducting sentiment analysis on dutch text to distinguish between positive and negative relations. *Journal of Information Technology and Politics, 5*(1), 73–94. https://doi.org/10.1080/19331680802154145

55. Young, L., & Soroka, S. (2012). Affective news: The automated coding of sentiment in political texts. *Political Communication, 29*(2), 205–231. https://doi.org/10.1080/10584609.2012.671234

56. Krawczyk, K., Chelkowski, T., Laydon, D. J., Mishra, S., Xifara, D., Gibert, B., Flaxman, S., Mellan, T., Schwämmle, V., Röttger, R., Hadsund, J. T., & Bhatt, S. (2021). Quantifying online news media coverage of the COVID-19 pandemic: Text mining study and resource. *Journal of Medical Internet Research*. https://doi.org/10.2196/28253

57. Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., & Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities and Social Sciences Communications, 7*(1), 1–23. https://doi.org/10.1057/s41599-020-0523-3

58. Catelli, R., Pelosi, S., Comito, C., Pizzuti, C., & Esposito, M. (2023). Lexicon-based sentiment analysis to detect opinions and attitude towards COVID-19 vaccines on Twitter in Italy. *Computers in Biology and Medicine, 158*, 1–16. https://doi.org/10.1016/j.compbiomed.2023.106876

59. Yu, X., Zhong, C., Li, D., & Xu, W. (2020). Sentiment analysis for news and social media in COVID-19. In *Proceedings of the 6th ACM SIGSPATIAL international workshop on emergency management using GIS* (pp. 1–4). https://doi.org/10.1145/3423333.3431794

60. Boomgaarden, H. G., & Schmitt-Beck, R. (2019). The media and political behavior. In H. G. Boomgaarden & R. Schmitt-Beck (Eds.), *Oxford research encyclopedia of politics.* Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.621

61. Ghasiya, P., & Okamura, K. (2021). Investigating COVID-19 news across four nations: A topic modeling and sentiment analysis approach. *IEEE Access, 9*, 36645–36656. https://doi.org/10.1109/ACCESS.2021.3062875

62. Robertson, F., Lagus, J., & Kajava, K. (2021). A COVID-19 news coverage mood map of Europe. *Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation*, 110–115. https://aclanthology.org/2021.hackashop-1.15

63. Pandur, M. B., Dobša, J., Beliga, S., & Meštrović, A. (2020, October 5). Topic modelling and sentiment analysis of COVID-19 related news on Croatian Internet portal. *INFORMATION SOCIETY 2020.* https://ailab.ijs.si/dunja/SiKDD2021/Papers/Panduretal.pdf

64. Ali, G. GMd. N., Rahman, Md. M., Hossain, Md. A., Rahman, Md. S., Paul, K. C., Thill, J.-C., & Samuel, J. (2021). Public perceptions of COVID-19 vaccines: Policy implications from US spatiotemporal sentiment analytics. *Healthcare, 9*(9), 1–32. https://doi.org/10.3390/healthcare9091110

65. Yin, H., Song, X., Yang, S., & Li, J. (2022). Sentiment analysis and topic modeling for COVID-19 vaccine discussions. *World Wide Web, 25*(3), 1067–1083. https://doi.org/10.1007/s11280-022-01029-y

66. Sear, R. F., Velasquez, N., Leahy, R., Restrepo, N. J., Oud, S. E., Gabriel, N., Lupu, Y., & Johnson, N. F. (2020). Quantifying COVID-19 content in the online health opinion war using machine learning. *IEEE Access, 8*, 91886–91893. https://doi.org/10.1109/ACCESS.2020.2993967

67. Zhou, A., Liu, W., & Yang, A. (2023). Politicization of science in COVID-19 vaccine communication: Comparing US politicians, medical experts, and government agencies. *Political Communication.* https://doi.org/10.1080/10584609.2023.2201184

68. Béres, F., Csoma, R., Michaletzky, T., & Benczúr, A. (2022). COVID vaccine sentiment dashboard based on Twitter data. *Scientia et Securitas, 2*(4), 418–427. https://doi.org/10.1556/112.2021.00066

69. Szabó, G. (Ed.). (2022). *Érzelmek és Járványpolitizálás Politikai érzelemmenedzserek és érzelemszabályozási ajánlataik Magyarországon a COVID-19 pandémia idején* (2022nd ed.). ELKH Társadalomtudományi Kutatóközpont.

70. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding.* arXiv:1810.04805, 4171–4186. https://doi.org/10.48550/arXiv.1810.04805

71. Gupta, S., Bolden, S. E., Kachhadia, J., Korsunska, A., & Stromer-Galley, J. (2020). PoliBERT: Classifying political social media messages with BERT. *Social, Cultural and Behavioral Modeling (SBP-BRIMS 2020) Conference.* http://sbp-brims.org/2020/proceedings/papers/working-papers/SBP-BRiMS_2020_paper_59.pdf

72. Kis, G. M., Ring, O., & Sebők, M. (2022). A novel cost-efficient use of BERT embeddings in 8-way emotion classification on a Hungarian Media Corpus [Preprint]. SocArXiv. https://osf.io/gwytr

73. Nemeskey, D. M. (2020). *Natural language processing methods for language modeling.* Eötvös Loránd University.

74. Sebők, M., Máté, Á., Ring, O., Kovács, V., & Lehoczki, R. (2024). Leveraging open large language models for multilingual policy topic classification: the babel machine approach. *Social Science Computer Review,* 0(0), 1–23. https://doi.org/10.1177/08944393241259434

75. Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). *GoEmotions: A dataset of fine-grained emotions.* Preprint arXiv:2005.00547. https://doi.org/10.48550/ARXIV.2005.00547

76. Koljonen, J., Öhman, E., Ahonen, P., & Mattila, M. (2022). Strategic sentiments and emotions in post-Second World War party manifestos in Finland. *Journal of Computational Social Science, 5*(2), 1529–1554. https://doi.org/10.1007/s42001-022-00181-9

77. Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*(5), 360–363.

78. Vliegenthart, R., Walgrave, S., Baumgartner, F. R., Bevan, S., Breunig, C., Brouard, S., Bonafont, L. C., Grossman, E., Jennings, W., Mortensen, P. B., Palau, A. M., Sciarini, P., & Tresch, A. (2016). Do the media set the parliamentary agenda? A comparative study in seven countries. *European Journal of Political Research, 55*(2), 283–301. https://doi.org/10.1111/1475-6765.12134

79. Froio, C., Bevan, S., & Jennings, W. (2017). Party mandates and the politics of attention: Party platforms, public priorities and the policy agenda in Britain. *Party Politics, 23*(6), 692–703. https://doi.org/10.1177/1354068815625228

80. Lovett, J., Bevan, S., & Baumgartner, F. R. (2015). Popular presidents can affect congressional attention, for a little while: Presidents can affect congressional attention. *Policy Studies Journal, 43*(1), 22–43. https://doi.org/10.1111/psj.12087

Authors and Affiliations

Miklós Sebők[1] · **Orsolya Ring[1]** · **Márk György Kis[1]** ·
Martin Balázs Bánóczy[1] · **Ágnes Dinnyés[1]**

✉ Miklós Sebők
sebok.miklos@tk.hun-ren.hu
https://scholar.google.com/citations?user=79Cx-j8AAAAJ

Orsolya Ring
https://scholar.google.com/citations?user=hvWHCC8AAAAJ

Márk György Kis
https://scholar.google.com/citations?user=tlJ-kYsAAAAJ

[1]   HUN-REN Centre for Social Sciences, Budapest, Hungary